Journal of Combinatorial Theory, Series A **83**, 233–250 (1998)
Article No. TA982876

Optimal Linear Perfect Hash Families

Simon R. Blackburn* and Peter R. Wild

Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom

Communicated by the Managing Editors

Received April 29, 1997

Let V be a set of order n and let F be a set of order q. A set $S \subseteq \{\phi: V \to F\}$ of functions from V to F is an (n, q, t)-*perfect hash family* if for all $X \subseteq V$ with $|X| = t$, there exists $\phi \in S$ which is injective when restricted to X. Perfect hash families arise

previously known lower bounds for many parameter sets. The paper exhibits new classes of perfect hash families which show that these lower bounds are realistic.

1. INTRODUCTION

Let n and q be integers such that $2 \leqslant q \leqslant n$. Let V be a set of order n and let F be a set of order q. For any subset $P \subseteq V$ and any function $\phi: V \to F$, we say that ϕ *separates* P if ϕ is injective when restricted to P. Let t be an integer such that $2 \leqslant t \leqslant q$, and let $S \subseteq \{\phi: V \to F\}$. We say that S is *an* (n, q, t)-*perfect hash family* if for all $P \subseteq V$ with $|P| = t$ there exists $\phi \in S$ such that ϕ separates P.

Perfect hash families first arose as part of database management—see Mehlhorn [12] for a summary of early results. They are involved in certain circuit complexity problems (see Newman and Wigderson [13]) and have been used in the design of deterministic analogues of probabilistic algorithms (see Alon and Naor [2]). There have also been recent cryptographic applications (see Blackburn, Burmester, Desmedt and Wild [4]; the connection with perfect hash families was pointed out by Kurosawa and Stinson [11]).

This paper addresses the following question: How small can a set S be, subject to being an (n, q, t)-perfect hash family? Answers to this question have implications for the formula size of certain Boolean functions [13],

* This author is an E.P.S.R.C. Advanced Fellow

0097-3165/98 $25.00

and provide information on the share expansion of certain secret sharing schemes which are relevant in threshold cryptography [4]. We provide a new lower bound on the cardinality of an (n, q, t)-perfect hash family S, and construct (using a linear construction) classes of perfect hash families which show that our lower bound is realistic. We also show that our construction is optimal amongst 'linear' perfect hash families. We discuss our results in more detail in the next two paragraphs.

In Section 2, we prove the following. Let S be an (n, q, t)-perfect hash family. Then, if $t=2$ and $n>q^e$ or $t\geqslant 3$ and $n>(t-1)(q^e-1)$, we have $|S|>(t-1)e$. Further, if $n>q^{e+1}/(t-1)+t(q^e-1)+q-1$ then $|S|>(t-1)e+1$.

Set $d=\log n/\log q$ (so $n=q^d$). With $e=\lceil d\rceil-1$, these bounds may be regarded as a generalisation and strengthening of the well known 'volume' bound [12], which states that $|S|\geqslant\lceil d\rceil$. Our bounds improve previously known bounds due to Körner and Marton [10] (which generalise bounds due to Fredman and Komlós [6]) for many parameter sets; see Section 2 for a comparison of our bounds with theirs.

An elementary probabilistic argument [12] implies that a (q^d, q, t)-perfect hash family S exists with $|S|=s$ whenever

$$s\geqslant\frac{\log\binom{q^d}{t}}{t\log q-\log\left(q^t-t!\binom{q}{t}\right)}.$$

Since the right hand side of this inequality tends to dt as $q\to\infty$ with d and t fixed, this argument implies that our lower bound is reasonable for many parameter sets. However, we are able to find classes of perfect hash families that have smaller cardinality than those produced by the straightforward probabilistic approach. Sections 3 and 4 consider the class of linear perfect hash families: A perfect hash family $S\subseteq\{\phi: V\to F\}$ is *linear* if F may be identified with a field and V with a vector space over F in such a way that S becomes a set of linear functionals. Of course, q must necessarily be a prime power and n must be an integer power of q. We present a probabilistic argument over all linear perfect hash families which implies that there exists a (linear) (q^d, q, t)-perfect hash family S with $|S|=d(t-1)$ for all sufficiently large prime powers q, thus bettering the straightforward probabilistic approach. We also show that no linear (q^d, q, t)-perfect hash family S exists such that $|S|<d(t-1)$. We call a linear (q^d, q, t)-perfect hash family *optimal* if $|S|=d(t-1)$.

A disadvantage of probabilistic approaches is that no specific classes of perfect hash families are constructed. However, our linear approach does produce explicit classes of optimal linear perfect hash families for certain

fields. These explicit families better previously known constructions, such as the construction from error correcting codes by Alon [1], Brickell's construction [5] using resolvable balanced incomplete block designs and the various inductive constructions of Atici, Magliveras, Stinson and Wei [3], for the parameters that we are concerned with. To the best of our knowledge, the explicit linear perfect hash families in this paper are the first constructions of perfect hash families that are of smaller cardinality than the families that probabilistic methods produce, other than the class constructed when $q = t = 3$ by Körner and Marton [10, Section 3].

We let $Y(n, q, t)$ denote the minimum of $|S|$ over all (n, q, t)-perfect hash families S. It is well known that Mehlhorn's volume bound can be met for $t = 2$ and all n and q. Thus $Y(n, q, 2) = \lceil d \rceil$. Our results determine $Y(n, q, 3)$ for many values of n and q. The existence of optimal linear (n, q, t)-perfect hash families together with our lower bounds means that $Y(n, q, 3) = 2\lceil d \rceil$ for many values of n and q. We also give a construction that shows that $Y(n, q, 3) = 3$ for many values of n and q with $\lceil d \rceil = 2$.

2. LOWER BOUNDS

Let n, q and t be integers such that $2 \leqslant t \leqslant q \leqslant n$. In this section we establish lower bounds on the size of an (n, q, t)-perfect hash family. The following lemma and theorem provide a generalisation of Mehlhorn's [12] volume bound (case (i) of Theorem 1 below).

Let V be a set of order n and let F be a set of order q. For any subset $R \subseteq \{\phi: V \to F\}$ and any subset $W \subseteq V$, let W_R denote the subset of W consisting of those elements $w \in W$ such that for every $v \in W \backslash \{w\}$ the subset $\{w, v\}$ is separated by some $\phi \in R$.

LEMMA 1. *Let $R \subseteq \{\phi: V \to F\}$ and let $W \subseteq V$. If $|W| > q^{|R|}$ then $|W_R| \leqslant q^{|R|} - 1$.*

Proof. Let $|R| = l$ and write $R = \{\phi_1, ..., \phi_l\}$. Suppose that $|W| > q^l$. For any $v \in V$ put $\underline{v} = (\phi_1(v), ..., \phi_l(v)) \in F^l$. Now $w \in W_R$ if and only if $\underline{w} \neq \underline{v}$ for every $v \in W \backslash \{w\}$. Thus the set $\{\underline{w} \in F^l : w \in W_R\}$ is a set of $|W_R|$ distinct elements of F^l. Also, since $|W| > q^l$, not every $(a_1, ..., a_l) \in F^l$ equals $\underline{v}$ for some $v \in W$. Hence $|W_R| < |F^l| = q^l$. ∎

THEOREM 1. *Let $S \subseteq \{\phi: V \to F\}$ be an (n, q, t)-perfect hash family. Let e be a positive integer.*

(i) *If $t = 2$ and $n > q^e$ then $|S| > e$;*

(ii) *If $t \geqslant 3$ and $n > (t-1)(q^e - 1)$ then $|S| > (t-1)e$.*

Proof. Suppose that $n > q^e$ if $t = 2$ or $n > (t-1)(q^e - 1)$ if $t \geqslant 3$ and suppose that $|S| \leqslant (t-1)e$. Let $R \subseteq \{\phi: V \to F\}$ be such that $S \subseteq R$ and $|R| = (t-1)e$. We show that there is a subset $P \subseteq V$ with $|P| = t$ such that P is not separated by any $\phi \in R$. It follows that S is not a perfect hash family and this contradiction proves the theorem.

Let $R = \{\phi_1, ..., \phi_{(t-1)e}\}$ and put $R_i = \{\phi_{(i-1)e+1}, ..., \phi_{ie}\}$ for $i = 1, ..., t-1$. Now $n > q^e$ since $(t-1)(q^e - 1) \geqslant q^e$ for $t \geqslant 3$. Hence, by Lemma 1, $|\bigcup_{i=1}^{t-1} V_{R_i}| \leqslant (t-1)(q^e - 1)$. Therefore there exists an element $v_t \in V \backslash \bigcup_{i=1}^{t-1} V_{R_i}$. For $i = 1, ..., t-1$ let $v_i \in V$ be such that $\{v_t, v_i\}$ is not separated by any $\phi \in R_i$. Such a v_i exists since $v_t \notin V_{R_i}$ for $i = 1, ..., t-1$. Let P be any subset with $|P| = t$ that contains $\{v_1, ..., v_t\}$. Then P is not separated by any $\phi \in R = \bigcup_{i=1}^{t-1} R_i$. ∎

COROLLARY 1. *Let t be an integer and $\hat{d}$ a real number such that $t \geqslant 2$ and $\hat{d} \geqslant 1$. Then, for all sufficiently large integers q, an (n, q, t)-perfect hash family $S \subseteq \{\phi: V \to F\}$ with $n \geqslant q^{\hat{d}}$ satisfies $|S| \geqslant (t-1)(\lceil \hat{d} \rceil - 1) + 1$.*

Proof. If $t = 2$ then $n \geqslant q^{\hat{d}} > q^{\lceil \hat{d} \rceil - 1}$. Hence $|S| \geqslant \lceil \hat{d} \rceil$. If $t \geqslant 3$ then, for sufficiently large q, we have $n \geqslant q^{\hat{d}} > (t-1)(q^{\lceil \hat{d} \rceil - 1} - 1)$. Hence $|S| \geqslant (t-1)(\lceil \hat{d} \rceil - 1) + 1$. ∎

With $e = \lceil d \rceil - 1$ where $d = \log n / \log q$, Theorem 1 shows that if n exceeds an integer power of q by an appropriate amount then $|S|$ is at least $(t-1)(\lceil d \rceil - 1) + 1$. Corollary 1 shows that, if t and d are fixed then, asymptotically in q, this bound always applies. The following theorem shows that, for $t \geqslant 3$, we can improve the lower bound by 1 when n is sufficiently close to $q^{\lceil d \rceil}$.

THEOREM 2. *Let $S \subseteq \{\phi: V \to F\}$ be an (n, q, t)-perfect hash family. If $n > q^{e+1}/(t-1) + t(q^e - 1) + q - 1$ then $|S| > (t-1)e + 1$.*

Proof. When $t = 2$ the bound follows from Mehlhorn's [12] volume bound.

Suppose that $t \geqslant 3$. Suppose that $n > q^{e+1}/(t-1) + t(q^e - 1) + q - 1$ and $|S| \leqslant (t-1)e + 1$. Let $R \subseteq \{\phi: V \to F\}$ be such that $S \subseteq R$ and $|R| = (t-1)e + 1$. We show that there is a subset $P \subseteq V$ with $|P| = t$ such that P is not separated by any $\phi \in R$. It follows that S is not a perfect hash family and this contradiction proves the theorem.

Let $R = \{\phi_0, ..., \phi_{(t-1)e}\}$. Put $R_0 = \{\phi_0\}$ and, for $i = 1, ..., t-1$, put $R_i = \{\phi_{(i-1)e+1}, ..., \phi_{ie}\}$. Let $W = V \backslash (\bigcup_{i=1}^{t-1} V_{R_i})$. By Lemma 1, $|W| > q^{e+1}/(t-1) + q^e + q - 2$.

For each $v \in V$, put $\underline{v}^i = (\phi_{(i-1)e+1}(v), ..., \phi_{ie}(v)) \in F^e$, for $i = 1, ..., t-1$. For $i = 1, ..., t-1$ define an equivalence relation $\equiv_i$ on V such that, for all $u, v \in V$, $u \equiv_i v$ if and only if $\underline{u}^i = \underline{v}^i$. Define an equivalence relation $\equiv_0$ on

W such that, for all $u, v \in W$, $u \equiv_0 v$ if and only if $\phi_0(u) = \phi_0(v)$. Let $C_1, ..., C_l$ denote the equivalence classes of the equivalence relation $\equiv_0$ that contain more than one element. Put $C = \bigcup_{i=1}^{l} C_i$. Then $l \leqslant q$ and $|C| > q^{e+1}/(t-1) + q^e - 1$.

Let $n_i^u = |[u]_i \cap C|$ where $[u]_i$ is the equivalence class of the equivalence relation $\equiv_i$ that contains $u \in V$. Also let $\{u_j : j \in J\}$ be a set of representatives of the equivalence classes of the equivalence relation $\equiv_i$. Now

$$\sum_{u \in C} n_i^u = \sum_{j \in J} (n_i^{u_j})^2 \geqslant \left(\sum_{j \in J} n_i^{u_j} \right)^2 \Big/ |J| = |C|^2/|J|.$$

Since there are at most q^e such classes we have

$$\sum_{u \in C} \left(\sum_{i=1}^{t-1} n_i^u \right) = \sum_{i=1}^{t-1} \left(\sum_{u \in C} n_i^u \right) \geqslant \sum_{i=1}^{t-1} |C|^2/q^e = (t-1)|C|^2/q^e.$$

So there exists an element $u \in C$ with

$$\sum_{i=1}^{t-1} n_i^u \geqslant (t-1)|C|/q^e > q + t - 2.$$

Suppose that for any distinct integers i, j with $1 \leqslant i, j \leqslant t-1$ we have $[u]_i \cap [u]_j = \{u\}$. If no two elements of $\bigcup_{k=1}^{t-1} [u]_k \cap C$ belonged to the same class C_i of the equivalence relation $\equiv_0$ then we would have

$$q \geqslant l \geqslant 1 + \sum_{i=1}^{t-1} (n_i^u - 1) = \left(\sum_{i=1}^{t-1} n_i^u \right) - t + 2.$$

But this is not satisfied for the element u. Hence either there exists an element $v \neq u$ belonging to $[u]_i$ for some i $(1 \leqslant i \leqslant t-1)$ such that $\{u, v\}$ is not separated by ϕ_0 or there exist elements $w_1, w_2 \in [u]_i \cap C$ for some i $(1 \leqslant i \leqslant t-1)$ such that $\{w_1, w_2\}$ is not separated by ϕ_0. In the first case put $u_t = u$, $u_i = v$ and in the second case put $u_t = w_1$ and $u_i = w_2$. Also, for $j = 1, ..., t-1$, $j \neq i$, let $u_j \in [u_t]_j \backslash \{u_t\}$. These elements exist since $u_t \in W$. Then the subset $\{u_1, ..., u_t\}$, and therefore any subset $P \subseteq V$ with $|P| = t$ that contains it, is not separated by any $\phi \in R$.

Suppose now that there exist distinct integers i, j with $1 \leqslant i, j \leqslant t-1$ such that there is an element $v \neq u$ belonging to $[u]_i \cap [u]_j$. Put $u_i = u$, $u_j = v$ and for $l = 1, ..., t-1$, $l \neq i, j$, let $u_l \in [u]_l \backslash \{u\}$. Finally, let u_t be such that u and u_t are distinct but belong to the same equivalence class C_k of the equivalence relation $\equiv_0$. We note that u_l $(1 \leqslant l \leqslant t, l \neq i, j)$ exist since $u \in W$. It is easily checked that $\{u_1, ..., u_t\}$, and therefore any subset $P \subseteq V$ with $|P| = t$ that contains it, is not separated by any $\phi \in R$. ▮

Theorems 1 and 2 give the following for $t = 3$.

COROLLARY 2. *Let $S \subseteq \{\phi: V \to F\}$ be an $(n, q, 3)$-perfect hash family. If $n > 2q^{\lceil d \rceil - 1}$ then $|S| \geqslant 2\lceil d \rceil - 1$ and if $n > q^{\lceil d \rceil}/2 + 2(q^{\lceil d \rceil - 1} - 1) + q - 1$ then $|S| \geqslant 2\lceil d \rceil$.*

In the next section we show that for for all integers $d \geqslant 2$ and sufficiently large prime powers q there exists a (linear) (n, q, t)-perfect hash family with $n = q^d$ and $|S| = (t-1)d$. It follows that $Y(n, q, 3) = 2\lceil d \rceil$ for many values of n and q. The example that follows shows that we also have $Y(n, q, 3) = 2\lceil d \rceil - 1$ for many values of n and q with $\lceil d \rceil = 2$.

Let q_1 be a prime power and let $\mathbb{F}_{q_1^2}$ be the field with q_1^2 elements. We construct a $(q_1^3 - q_1, q_1^2 - 1, 3)$-perfect hash family S with $|S| = 3 = \lceil 2 \log(q_1^3 - q_1)/\log(q_1^2 - 1) \rceil$ from the classical unital in the Desarguesian projective plane of order q_1^2. This object (see, for example, Hughes and Piper [8]) consists of $q_1^3 + 1$ points and has the property that every line of the plane is either a secant and contains $q_1 + 1$ points of the unital or is a tangent and contains 1 point of the unital. Each point of the unital is on exactly one tangent.

Fix a secant to the unital. Let V be the set of points of the unital not belonging to this secant. Let L be a set of three points of the unital belonging to this secant. Each point of L is on $q_1^2 - 1$ other secants. Label these lines with the elements of a set F of size $q_1^2 - 1$. Then L determines in a natural way a set S of three functions $\phi: V \to F$ (the image of a point is the label of the line that joins it to the point of L that gives rise to the function ϕ).

Now if a set P of 3 points of V is not separated by any of the three functions in S then it determines a configuration of six points $P \cup L$ lying on four lines with each line containing three of the six points. Such a configuration is known as an O'Nan configuration. The classical unital is characterised by the property that it contains no O'Nan configurations (O'Nan [14]). It follows that every set P of 3 points of V is separated by some $\phi \in S$. Hence S is a $(q_1^3 - q_1, q_1^2 - 1, 3)$-perfect hash family with $|S| = 3$.

A result of Fredman and Komlós [6] (which has been restated in terms of graph entropy by Körner [9]) implies that

$$|S| \geqslant \frac{\binom{n-1}{t-2} q^{t-2} \log(n-t+2)}{\binom{q-1}{t-2} n^{t-2} \log(q-t+2)}.$$

Thus the bound is asymptotically equal to

$$\frac{q^{t-1}}{q(q-1)(q-2)\cdots(q-t+2)}\,\frac{\log n}{\log(q-t+2)}$$

as $n \to \infty$ with t fixed. (By two functions being asymptotically equal we mean that the ratio of the functions tends to 1.) This bound is not elementary to prove—see Radhakrishnan [15] for a slightly weaker bound with an elementary proof. Körner and Marton [10] have extended the entropy argument of [9] to establish a bound which is stronger than the Fredman–Komlós bound for many parameter sets; they show that $|S|$ is asymptotically bounded below by

$$\min_{0 \leqslant j \leqslant t-2} \frac{q^{j+1}}{q(q-1)(q-2)\cdots(q-j)}\,\frac{\log n}{\log\left(\dfrac{q-j}{t-j-1}\right)}.$$

The Fredman–Komlós and the Körner–Marton bounds are better than our bound when t is close to q. For example, if $q \to \infty$ with $d = \log n/\log q$ and $q-t$ fixed (where $d > 1$), then the Fredman–Komlós and Körner–Marton bounds are exponential in q, whereas our bound is linear in q. If $n \to \infty$ with q and t fixed, our bound is stronger for many values of t and q. Table I lists constants $\kappa_{q,t}$ where the best of the Mehlhorn, Fredman–Komlós, Marton–Körner, and our bounds, is asymptotically equal to $\kappa_{q,t}d$ as $n \to \infty$. (For sufficiently large n, our bound is stronger when $t = 3$ and $q \geqslant 5$, when $t = 4$ and $q \geqslant 8$, when $t = 5$ and $q \geqslant 12$ or when $t = 6$ and $q \geqslant 16$ for example).

TABLE I

The Constants $\kappa_{q,t}$

$q\backslash t$	2	3	4	5	6	7	8
2	1.000						
3	1.000	3.566					
4	1.000	2.243	16.000				
5	1.000	2.000	6.782	90.700			
6	1.000	2.000	4.362	26.095	628.126		
7	1.000	2.000	3.318	13.165	122.507	5160.664	
8	1.000	2.000	3.000	8.467	48.000	680.636	49152.000
9	1.000	2.000	3.000	6.206	25.797	205.643	4374.000
10	1.000	2.000	3.000	4.928	16.526	91.563	1013.772
11	1.000	2.000	3.000	4.122	11.808	50.816	371.633
12	1.000	2.000	3.000	4.000	9.074	32.411	177.517

When $q \to \infty$ with t and $d = \log n/\log q$ fixed, the Fredman–Komlós and Körner–Marton bounds are asymptotically equal to d, hence our bound is always asymptotically stronger in this case. Thus our bound improves upon previously known bounds for many parameter sets.

The basis of the Fredman–Komlós and Körner–Marton bounds is a counting argument used in conjunction with a volume bound. Radhakrishnan [15] has given a weaker bound based on a simpler counting argument. We now show that Radhakrishnan's argument can be extended in the same way that Körner and Marton [10] extended the argument of Fredman and Komlós [6] and can be used in conjunction with bounds other than the volume bound.

Let $S \subseteq \{\phi: V \to F\}$ be an (n, q, t)-perfect hash family. Let $Q \subseteq V$ be a set of size j where $0 \leqslant j \leqslant t-2$. Let S_Q be the subset of S consisting of those functions that separate Q. Let $\bar{V} = V \backslash Q$ and let $\bar{F}$ be a set of size $q-j$. For each $\phi \in S_Q$ we define a function $\bar{\phi}: \bar{V} \to \bar{F}$ as follows. Let π be an arbitrary bijection $\pi: F \backslash \phi(Q) \to \bar{F}$ (since $\phi \in S_Q$ separates Q we have $|F \backslash \phi(Q)| = |\bar{F}|$). Let α be a fixed element in $\bar{F}$. For each $v \in \bar{V}$ put $\bar{\phi}(v) = \alpha$ if $\phi(v) \in \phi(Q)$ and $\bar{\phi}(v) = \pi(\phi(v))$ otherwise. Finally we define $\bar{S}_Q = \{\bar{\phi}: \phi \in S_Q\}$.

PROPOSITION 1. *Let $S \subseteq \{\phi: V \to F\}$ be an (n, q, t)-perfect hash family. Let Q be a subset of V of size j where $0 \leqslant j \leqslant t-2$. Then $\bar{S}_Q$ defined above is an $(n-j, q-j, t-j)$-perfect hash family.*

Proof. Clearly $|\bar{V}| = n-j$ and $|\bar{F}| = q-j$. Now let $P \subseteq \bar{V}$ be a set of size $t-j$. Then $Q \cup P$ is a subset of V of size t. Hence there exists a $\phi \in S$ that separates $Q \cup P$. As ϕ separates Q, we have $\phi \in S_Q$. Since ϕ restricted to $Q \cup P$ is injective, $\phi(v) \notin \phi(Q)$ for each $v \in P$, and so $\bar{\phi}$ is injective when restricted to P. Thus $\bar{\phi}$ separates P. It follows that $\bar{S}$ is an $(n-j, q-j, t-j)$-perfect hash family. ∎

COROLLARY 3. *Let $S \subseteq \{\phi: V \to F\}$ be an (n, q, t)-perfect hash family. The number of functions $\phi \in S$ that separate a given j-subset of V $(0 \leqslant j \leqslant t-2)$ is at least $Y(n-j, q-j, t-j)$.*

The following result is an easy extension of the counting argument of Radhakrishnan [15].

THEOREM 3. *Let $S \subseteq \{\phi: V \to F\}$ be an (n, q, t)-perfect hash family. Then for $0 \leqslant j \leqslant t-2$ we have*

$$|S| \geqslant \frac{q^j \binom{n}{j}}{n^j \binom{q}{j}} Y(n-j, q-j, t-j).$$

Proof. We count pairs (P, ϕ) where P is a j-subset of V and $\phi \in S$ separates P. There are $\binom{n}{j}$ subsets P and, by Corollary 3, each is separated by at least $Y(n-j, q-j, t-j)$ functions $\phi \in S$. Also each function in S separates at most $\binom{q}{j}\left(\frac{n}{q}\right)^j$ subsets of size j. Hence

$$|S| \left(\frac{n}{q}\right)^j \binom{q}{j} \geqslant \binom{n}{j} Y(n-j, q-j, t-j)$$

and the result follows. ∎

COROLLARY 4. *Let $S \subseteq \{\phi: V \to F\}$ be an (n, q, t)-perfect hash family. Then*

$$Y(n, q, t) \geqslant \max_{0 \leqslant j \leqslant t-2} \left\{ \frac{q^j \binom{n}{j}}{n^j \binom{q}{j}} Y(n-j, q-j, t-j) \right\}.$$

Corollary 4 may be used to derive bounds for a given set of parameters from bounds for smaller parameter values. It is possible to combine Corollary 4 with our bounds. However our investigations suggest that, asymptotically as $n \to \infty$, there is a boundary in the plane determined by the parameters q and t, one side of which the Fredman–Komlós bound is superior and the other side of which our bound is superior without the benefit of Corollary 4.

3. LINEAR PERFECT HASH FAMILIES

Let $S \subseteq \{\phi: V \to F\}$ be an (n, q, t)-perfect hash family. We say that S is a *linear perfect hash family* if F can be identified with a finite field $\mathbb{F}_q$ and V can be identified with a vector space over $\mathbb{F}_q$ in such a way that S is a set of linear functionals under this identification. For S to be linear, it is of course necessary for q to be a prime power and for $n = q^d$ for some non-negative integer d. This section contains the statements of the two theorems we will prove concerning linear perfect hash families, and establishes the notation that we use in the proof of these theorems in Section 4.

THEOREM 4. *Let d and t be integers such that $d \geqslant 2$ and $t \geqslant 2$ and let q be a prime power. Set $n = q^d$. If S is a linear (n, q, t)-perfect hash family, then $|S| \geqslant d(t-1)$. Furthermore, if $q \geqslant (\frac{1}{2}t(t-1))^{d(t-1)}$ then a linear (n, q, t)-perfect hash family S exists with $|S| = d(t-1)$.*

The proof of Theorem 4 is probabilistic, and so produces no explicit classes of perfect hash families. However, the techniques of Theorem 4 suffice to construct explicit classes of optimal linear perfect hash families in certain fields. In Section 4, we will prove the following theorem:

THEOREM 5. *Let d and t be integers such that $d \geqslant 2$ and $t \geqslant 2$ and let q be a prime power. Suppose that there exist finite fields $F_0 < F_1 < \cdots < F_{d(t-2)}$ such that $|F_{d(t-2)}| = q$ and such that $[F_i : F_{i-1}] \geqslant d$ for all integers $i \in \{1, 2, ..., d(t-2)\}$.*

Define a sequence $(\alpha^1, \alpha^2, ..., \alpha^{d(t-1)})$ of row vectors of length d as follows. For all integers i such that $1 \leqslant i \leqslant d$, define α^i to be the ith standard basis vector. For all integers i such that $d+1 \leqslant i \leqslant d(t-1)$, define

$$\alpha^i = (\beta_1^i, \beta_2^i, ..., \beta_d^i)$$

where $\{\beta_1^i, \beta_2^i, ..., \beta_d^i\}$ is any subset of F_{i-d} which is linearly independent over F_{i-d-1}. Set $V = (F_{d(t-2)})^d$ and define functionals $\phi_1, \phi_2, ..., \phi_{d(t-1)}$ by

$$(v)\phi_i = v(\alpha^i)^T$$

for all $v \in V$. Then $S = \{\phi_1, \phi_2, ..., \phi_{d(t-1)}\}$ is an optimal linear (q^d, q, t)-perfect hash family.

We now define the notation that we will use throughout this section and the next. As some of the notation is somewhat technical, we add some motivation in parentheses. Any parts of the motivational statements that we need will be proved during the proof of Theorem 5.

Let V be a vector space of dimension d over $\mathbb{F}_q$ and let V^* be the dual space of V, consisting of the set of linear functionals $\phi: V \to \mathbb{F}_q$. Given a set $S = \{\phi_1, ..., \phi_k\}$ of linear functionals we may order them in some arbitrary way to produce a sequence $(\phi_1, ..., \phi_k) \in (V^*)^k$. Our aim is to find a condition that $(\phi_1, ..., \phi_k)$ must satisfy in order to guarantee that S be a perfect hash family.

Let X be the dt-dimensional vector space over $\mathbb{F}_q$ defined by $X = (V^*)^t$. Let $P = \{p_1, p_2, ..., p_t\} \subseteq V$. Then P gives rise to a subspace U_P of X given by

$$U_P = \left\{ (\psi_1, \psi_2, ..., \psi_t) \in X : \sum_{1 \leqslant i \leqslant t} (p_i)\psi_i = 0 \right\}.$$

Let $T = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} : 1 \leqslant a < b \leqslant t\}$. For all $(a, b) \in T$ and all $\phi \in V^*$, define the vector $w_{(a,b),\phi} \in X$ by $w_{(a,b),\phi} = (\psi_1, \psi_2, ..., \psi_t)$ where

$$\psi_i = \begin{cases} \phi & \text{if } i = a \\ -\phi & \text{if } i = b \text{ and} \\ 0 & \text{if } i \neq a, b. \end{cases}$$

(Let $P = \{p_1, ..., p_t\} \subseteq V$ be a set of t distinct elements of V and let ϕ be a linear functional. One may show that ϕ fails to separate P if and only if $w_{(a,b),\phi} \in U_P$ for some $(a, b) \in T$.)

For all $(a, b) \in T$, define the subspace $V_{(a,b)}$ of X by

$$V_{(a,b)} = \{w_{(a,b),\phi} : \phi \in V^*\}.$$

(The set $P \subseteq V$ is of cardinality strictly less than t if and only if $V_{(a,b)} \subseteq U_P$ for some $(a, b) \in T$.)

Let $c = ((a_1, b_1), ..., (a_k, b_k)) \in T^k$ and let $\Phi = (\phi_1, ..., \phi_k) \in (V^*)^k$. Define $W_{c,\Phi}$ to be the subspace of X given by

$$W_{c,\Phi} = \langle w_{(a_i,b_i),\phi_i} : 1 \leqslant i \leqslant k \rangle.$$

(Let P be a set of t distinct elements of V and let S be the set of linear functionals corresponding to Φ. Then S fails to separate P if and only if there exists $c \in T^k$ such that $W_{c,\Phi} \subseteq U_P$.)

4. PROOFS OF THEOREMS 4 AND 5

We begin this section by proving a series of three technical results, and then prove Theorems 4 and 5.

LEMMA 2. *Let d be an integer such that $d \geqslant 2$. Let V be a vector space of dimension d over the finite field $\mathbb{F}_q$ of order q. Let $U_1, ..., U_m$ be m proper subspaces of V. Let $K = \bigcup_{i=1}^{m} U_i$. If $q \geqslant m$, then $K \neq V$.*

Proof. We have that

$$|K| = \left| \{0\} \cup \left(\bigcup_{i=1}^{m} (U_i \backslash \{0\}) \right) \right| \leqslant 1 + m(q^{d-1} - 1) < q^d.$$

Hence $K \neq V$. ∎

We will now use Lemma 2 to prove the main technical result of this section. Throughout this section, let d and t be integers such that $d, t \geqslant 2$, let q be a prime power and let V be a vector space of dimension d over $\mathbb{F}_q$. Define the set T, the vector spaces X, $V_{(a,b)}$ and $W_{c,\Phi}$ and the vectors $w_{(a,b),\phi}$ as in Section 3.

LEMMA 3. *Using the above notation, let k be an integer such that $1 \leqslant k \leqslant d(t-1)$. Suppose that $q \geqslant (\frac{1}{2}t(t-1))^{d(t-1)}$. Then there exists $\Phi \in (V^*)^k$ with the following property. For all $c \in T^k$ either*

(i) *the subspace $W_{c,\Phi}$ has dimension k, or*

(ii) *there exist $(a, b) \in T$ such that $V_{(a,b)} \subseteq W_{c,\Phi}$.*

Proof. We will prove the lemma by induction on k. When $k = 1$, any $\phi \in V^* \backslash \{0\}$ will do, for then the vector $w_{a_1, b_1, \phi}$ is non-zero for any distinct $a_1, b_1 \in \{1, ..., t\}$, and hence Case (i) always holds.

Suppose that $k > 1$ and assume, as an inductive hypothesis, that lemma holds for all smaller values of k. Let $\Phi' = (\phi_1, ..., \phi_{k-1}) \in (V^*)^{k-1}$ be such that for all $c' \in T^{k-1}$ either $W_{c',\Phi'}$ has dimension $k-1$ or there exists $(a, b) \in T$ such that $V_{(a,b)} \subseteq W_{c',\Phi'}$. Such a Φ' exists, by our inductive hypothesis. Let $E \subseteq T^k$ be defined by

$$E = \{(c', (a_k, b_k)) \in T^{k-1} \times T : \text{for all } (a, b) \in T,\ V_{(a,b)} \not\subseteq W_{c',\Phi'}\}.$$

Note that $|E| \leqslant T^k = (\frac{1}{2}t(t-1))^k$.

Let $e = (c', (a_k, b_k)) \in E$. Then we may associate a subspace $U_e \subseteq V^*$ with $e \in E$ by defining

$$U_e = \{\phi \in V^* : w_{(a_k, b_k), \phi} \in W_{c',\Phi'}\}.$$

Note that by the definition of E, U_e is always a proper subspace of V^*. Let $\phi_k \in V^* \backslash \bigcup_{e \in E} U_e$. Such a functional exists by Lemma 2, since we have that $|E| \leqslant (\frac{1}{2}t(t-1))^k \leqslant (\frac{1}{2}t(t-1))^{d(t-1)} \leqslant q$. Define $\Phi = (\phi_1, ..., \phi_k)$. We aim to show that Φ satisfies the conditions of the lemma.

Let $e = (c', (a_k, b_k)) \in T^k$. If $e \notin E$, then $W_{c,\Phi}$ falls under Case (ii) of the lemma, since $W_{e,\Phi}$ contains the subspace $W_{c',\Phi'}$. If $e \in E$, then $W_{c',\Phi'}$ has dimension $k-1$, by our choice of Φ'. Furthermore, our choice of ϕ_k implies that $w_{a_k, b_k, \phi_k} \notin W_{c',\Phi'}$. This implies that $W_{e,\Phi}$ has dimension k, and so Case (i) of the lemma is satisfied. We have now shown that for our choice of Φ, for all $e \in T^k$ either Case (i) or Case (ii) of the lemma holds. Hence Φ satisfies the conditions of the lemma, and so the result follows by induction on k. ∎

LEMMA 4. *Suppose that $q \geqslant (\frac{1}{2}t(t-1))^{d(t-1)}$. Let $k = d(t-1)$. Then there exists $\Phi \in (V^*)^k$ with the following property: For all $c \in T^k$, there exists $(a, b) \in T$ such that $V_{(a,b)} \subseteq W_{c,\Phi}$.*

Proof. Let Φ be chosen to satisfy the conditions of Lemma 3. Suppose, for a contradiction, that the Lemma 4 does not hold for this choice of Φ. Then our choice of Φ implies that $W_{c,\Phi}$ has dimension $r = d(t-1)$ for

some $c \in T^k$. It is clear that $W_{c, \Phi}$ is contained in the $d(t-1)$ dimensional subspace Y of X defined by

$$Y = \{(\psi_1, \psi_2, ..., \psi_t) \in X: \sum_{i=1}^{t} \psi_i = 0\}.$$

Now $\dim Y = \dim W_{c, \Phi}$, hence $W_{c, \Phi} = Y$. But Y contains $V_{(a, b)}$ for all $(a, b) \in T$. This contradiction proves the lemma. ▌

We are now in a position to prove Theorem 4:

Proof of Theorem 4. Let $S = \{\phi_1, \phi_2, ..., \phi_k\}$ be a set of linear functionals from V to $\mathbb{F}_q$, and suppose that $k < d(t-1)$. We construct a set $P \subseteq V$ of cardinality t which is not separated by S. This will show that S is not an (n, q, t)-perfect hash family, as required.

Without loss of generality, we may assume that $k = d(t-1) - 1$. We construct a sequence $p_1, ..., p_t$ of elements of V as follows. Let $p_1 \in V$ be chosen arbitrarily. Define

$$A_0 = \{v \in V: (v)\,\phi_i = (p_1)\,\phi_i \text{ for all } i \in \{1, 2, ..., d-1\}\}.$$

Then A_0, being the non-empty intersection of $d-1$ affine subspaces of dimension $d-1$ or d, is an affine subspace of V of dimension at least 1; thus, A_0 is strictly larger than $\{p_1\}$. Choose $p_2 \in A_0 \backslash \{p_1\}$. Let j be an integer such that $1 \leqslant j \leqslant t-2$. We define p_{j+2} as follows. Let

$$S_j = \{i \in \{dj, dj+1, ..., dj+d-1\}: (p_1)\phi_i \neq (p_2)\phi_i\}.$$

Define the affine subspace A_j of V by

$$A_j = \{v \in V: (v)\phi_i = (p_1)\phi_i \text{ for all } i \in S_j\}.$$

The method of choosing p_j splits into two cases:

Case I. A_j is strictly larger than $\{p_1\}$. Choose $p_j \in A_j \backslash \{p_1\}$. We have that $p_j \neq p_1$ in this case, but we do not exclude the possibility that $p_j = p_2$. In any case, for each i such that $dj \leqslant i \leqslant dj + d - 1$ either $(p_1)\phi_i = (p_2)\phi_i$ or $(p_1)\phi_i = (p_j)\phi_i$.

Case II. $A_j = \{p_1\}$. Since A_j is affine of dimension 0 and is the intersection of at most $|S_j|$ affine subspaces of dimension $d-1$, we must have $S_j = \{dj, dj+1, ..., dj+d-1\}$. In particular, $(p_1)\,\phi_{dj} \neq (p_2)\,\phi_{dj}$ and $(p_1)\,\phi_{dj+1} \neq (p_2)\,\phi_{dj+1}$. Define the affine subspace A'_j of V by

$$A'_j = \{v \in V: (v)\phi_i = (p_1)\phi_i \text{ for all } i \in S_j \backslash \{dj\}\}.$$

Now A_j' has dimension at least 1; indeed, A_j' has dimension exactly 1, since A_j is the intersection of A_j' with the affine hyperplane defined by

$$\{v \in V: (v)\,\phi_{dj} = (p_1)\,\phi_{dj}\}.$$

The function ϕ_{dj} cannot be constant when restricted to A_j', for then $A_j = A_j'$. So we may choose p_j to be the unique element of A_j' such that $(p_j)\,\phi_{dj} = (p_2)\,\phi_{dj}$. Note that $p_j \neq p_1$ as $(p_j)\,\phi_{dj} = (p_2)\,\phi_{dj} \neq (p_1)\,\phi_{dj}$. Furthermore, $p_j \neq p_2$, as $(p_j)\,\phi_{dj+1} = (p_1)\,\phi_{dj+1} \neq (p_2)\,\phi_{dj+1}$.

Not all the elements $p_1, p_2, ..., p_t$ are necessarily distinct. However, for each functional $\phi \in S$, there exist distinct elements p_a, p_b in this sequence such that $(p_a)\phi = (p_b)\phi$. Therefore, if we define P to be any set of t elements of V which contains the set $\{p_1, p_2, ..., p_t\}$, we find that P is not separated by S, as required.

Now suppose that $q \geqslant (\frac{1}{2}t(t-1))^{d(t-1)}$. We show that a linear (n, q, t)-perfect hash family S exists such that $|S| = d(t-1)$.

Let $k = d(t-1)$. Let $\Phi = (\phi_1, ..., \phi_k) \in (V^*)^k$ be such that (in the notation defined at the beginning of the section) for all $c \in T^k$, there exists $(a, b) \in T$ such that $V_{(a,b)} \subseteq W_{c,\Phi}$. Such a Φ exists, by Lemma 4. Set $S = \{\phi_1, \phi_2, ..., \phi_k\}$. We claim that S is a linear (n, q, t)-perfect hash family, i.e. that S separates any set P of t distinct elements of V.

Suppose, for a contradiction, that $P = \{p_1, p_2, ..., p_t\}$ is a set of t distinct elements of V that is not separated by S. For any integer i such that $1 \leqslant i \leqslant k$, ϕ_i does not separate P, so there exists $(a_i, b_i) \in T$ such that

$$(p_{a_i})\phi_i - (p_{b_i})\phi_i = 0. \tag{1}$$

Define $c \in T^k$ by $c = ((a_1, b_1), ..., (a_k, b_k))$.

The equalities (1) imply that $w_{(a_i, b_i), \phi_i} \in U_P$ for all i such that $1 \leqslant i \leqslant k$. Hence $W_{c,\Phi} \subseteq U_P$. By our choice of Φ, there exists $(a, b) \in T$ such that $V_{(a,b)} \subseteq W_{c,\Phi}$. Hence, by the definitions of $V_{(a,b)}$ and U_P, $(p_a)\phi - (p_b)\phi = 0$ for all $\phi \in V^*$. Therefore $p_a = p_b$. But $a \neq b$ and the elements $p_1, ..., p_t$ are distinct. This contradiction implies that no set P of size t fails to be separated by S. Hence S is a linear (n, q, t)-perfect hash family, as required. ∎

We will now prove Theorem 5. The explicit construction of 5 is based on the same techniques that are used in Theorem 4; both theorems use the fact that any sequence $\Phi \in (V^*)^{d(t-1)}$ satisfying the conditions of Lemma 3 is an optimal linear perfect hash family. However, rather than using a probabilistic construction to establish the existence of such a sequence, we show that the explicit sequence derived from the statement of Theorem 5 satisfies the conditions of Lemma 3.

Proof of Theorem 5. Let $\Phi = (\phi_1, \phi_2, ..., \phi_{d(t-1)})$, where the ϕ_i are defined as in the statement of Theorem 5. The proof of Lemma 4 and Theorem 4 implies that it is sufficient to show that for all $c \in T^{d(t-1)}$ either $\dim W_{c,\Phi} = d(t-1)$ or $V_{(a,b)} \subseteq W_{c,\Phi}$ for some $(a, b) \in T$. To this end, let $c = ((a_1, b_1), (a_2, b_2), ..., (a_{d(t-1)}, b_{d(t-1)})) \in T^{d(t-1)}$ and suppose that $\dim W_{c,\Phi} < d(t-1)$. Define k to be the smallest integer such that

$$\dim W_{((a_1, b_1), ..., (a_k, b_k)), (\phi_1, ..., \phi_k)} < k.$$

We show that $V_{(a_k, b_k)} \subseteq W_{c,\Phi}$.

First note that the linear independence of $\phi_1, \phi_2, ..., \phi_d$ implies that the vectors $w_{(a_1, b_1), \phi_1}, ...w_{(a_d, b_d), \phi_d}$ are linearly independent for all choices of $((a_1, b_1), ..., (a_d, b_d)) \in T^d$. Hence we must have that $k > d$.

Define $c' = ((a_1, b_1), ..., (a_{k-1}, b_{k-1}))$ and define $\Phi' = (\phi_1, \phi_2, ..., \phi_{k-1})$. Let the subspace $U \subseteq X$ be defined by

$$U = V_{(a_k, b_k)} \cap W_{c', \Phi'}.$$

By definition, $w_{(a_k, b_k), \phi_k} \in V_{(a_k, b_k)}$; by our choice of k we have that $w_{(a_k, b_k), \phi_k} \in W_{c', \Phi'}$. Thus $w_{(a_k, b_k), \phi_k} \in U$.

If $\dim U = d$, then $V_{(a_k, b_k)} = U \subseteq W_{c', \Phi'} \subseteq W_{c, \Phi}$ as required, so we assume that $\dim U < d$ and derive a contradiction. The standard basis of $V = (\mathbb{F}_q)^d$ gives rise to a basis $\{x_{i,j}: 1 \leqslant i \leqslant t, 1 \leqslant j \leqslant d\}$ of $X = (V^*)^t$. With respect to this basis, both $V_{(a,b)}$ and $W_{c', \Phi'}$ may be generated by vectors whose components are in the subfield F_{k-d-1}, since

$$V = \langle x_{a,i} - x_{b,i}: 1 \leqslant i \leqslant d \rangle \text{ and}$$

$$W_{c', \Phi'} = \left\langle w_{(a_i, b_i), \phi_i} = \sum_{j=1}^{d} (x_{a_i, j} - x_{b_i, j})\, \beta_j^i : 1 \leqslant j \leqslant k-1 \right\rangle.$$

Hence U, being the intersection of these subspaces, may be generated by vectors $u_1, u_2, ..., u_{d-1}$ having components in F_{k-d-1}. Since $w_{(a_k, b_k), \phi_k} \in U$, we have that there exist $\gamma_1, ..., \gamma_{d-1} \in F_{d(t-2)}$ such that

$$w_{(a_k, b_k), \phi_k} = \gamma_1 u_1 + \gamma_2 u_2 + \cdots + \gamma_{d-1} u_{d-1}. \tag{2}$$

Indeed, since all the vectors involved in (2) have components in F_{k-d}, we may take $\gamma_1, ..., \gamma_{d-1} \in F_{k-d}$. Comparing coefficients of $x_{a_k, i}$ in (2) shows that β_i^k is contained in the F_{k-d-1}-linear span of $\gamma_1, \gamma_2, ..., \gamma_{d-1}$, since the components of $u_1, u_2, ..., u_{d-1}$ are in F_{k-d-1}. But the elements $\beta_1^k, ..., \beta_d^k$ were chosen to be linearly independent over F_{k-d-1}, and therefore cannot

all be contained in the F_{k-d-1}-linear span of any $d-1$ elements of F_{k-d}. This contradiction completes the proof. ∎

5. DISCUSSION

The study of linear (n, q, t)-perfect hash families is of interest in Finite Geometry due to the following geometric interpretation. We may identify the elements of V with the points of the affine geometry $\mathrm{AG}(d, q)$ of dimension d over $\mathbb{F}_q$. For any linear functional $\phi: V \to \mathbb{F}_q$ and any element $\gamma \in \mathbb{F}_q$ the points $v \in V$ with $\phi(v) = \gamma$ form a hyperplane and so ϕ corresponds to a parallel class of hyperplanes. The property required of a set of parallel classes to determine a linear (n, q, t)-perfect hash family is that any t points of $\mathrm{AG}(d, q)$ should belong to distinct hyperplanes of some parallel class. Questions concerning the existence of linear (n, q, t)-perfect hash families relate to properties and existence of certain arcs in affine and projective space.

There is a correspondence between the parallel classes of $\mathrm{AG}(d, q)$ and the hyperplanes of the projective space $\mathrm{PG}(d-1, q)$ (the hyperplane at infinity of the projective space $\mathrm{PG}(d, q)$ that embeds $\mathrm{AG}(d, q)$). Two points of $\mathrm{AG}(d, q)$ belong to different hyperplanes of a parallel class if and only if the projective line joining them does not meet $\mathrm{PG}(d-1, q)$ in a point of the hyperplane of $\mathrm{PG}(d-1, q)$ corresponding to the parallel class. The proof of Theorem 4 shows that if S is a set of $(t-1)d$ linear functionals $\phi: V \to \mathbb{F}_q$ that contains a subset of d linear functionals which are linearly dependent then S is not an (n, q, t)-perfect hash family. Thus a necessary condition for a set of hyperplanes in $\mathrm{PG}(d-1, q)$ to correspond to an optimal linear (n, q, t)-perfect hash family is that they form a dual arc, that is no d of them are linearly dependent.

Theorem 4 provides a tight bound on the minimum cardinality of a linear (q^d, q, t)-perfect hash family, provided that the order of the underlying field is large enough. When the field order is below the bound of the theorem, it is expected that minimum cardinality of a linear (q^d, q, t)-perfect hash family will become strictly greater than $(t-1)d$: When does this begin to happen?

QUESTION 1. *Let t and d be integers such that $t, d \geqslant 2$. What is the largest prime power q such that no linear (q^d, q, t)-perfect hash family S exists with $|S| = (t-1)d$?*

Another way of phrasing this question is: Is the bound on the field order given in Theorem 4 reasonable?

There is a weaker upper bound on the number of functionals needed to separate every set of t points in a vector space, which holds for fields of much smaller order than Theorem 4. Indeed, the construction using dual arcs in [4, Section 4] (which is essentially Alon's coding-theoretic construction [1], using a doubly extended Reed–Solomon code) shows that whenever $q \geqslant \binom{t}{2}(d-1)$, there exists a linear (q^d, q, t)-perfect hash family S with $|S| = \binom{t}{2}(d-1)+1$.

When q is small compared to t, no linear (q^d, q, t)-perfect hash family can exist, for there exist subsets of a vector space of dimension d over $\mathbb{F}_q$ that are not separated by any linear functional. When $(q+4)/2 \leqslant t$ an example of such a subset is determined as follows. An oval of $\mathrm{PG}(2, q)$ is a set of $q+1$ points with the property that every line meets it in 0, 1 or 2 points. If q is odd then each point of the plane not belonging to the oval belongs to either $(q-1)/2$ or $(q+1)/2$ lines that join two points of the oval. Hence, for $q \geqslant 5$, any subset of $(q+3)/2$ points of an oval has the property that every point not on the oval belongs to at least one secant to the subset, *ie* a line meeting the subset in 2 points. Since we may take the oval to be contained in $\mathrm{AG}(2, q)$, such a subset determines a set P of $(q+3)/2$ points of a 2-dimensional subspace of V with the property that every parallel class of lines of that subspace contains a secant. Hence every parallel class of hyperplanes of V contains a hyperplane meeting P in at least 2 elements. When q is even, $q > 4$, we may show by a similar argument that there is a subset P of $(q+2)/2$ points of a 2-dimensional subspace of V with the same property.

QUESTION 2. *Let V be a vector space of dimension d over a finite field $\mathbb{F}_q$. What is the order of the smallest set P such that P cannot be separated by any linear functional*?

The above construction shows that this order can be at most $(q+4)/2$.

The lower bounds of Section 2 show that optimal linear perfect hash families have close to minimal cardinality amongst all perfect hash families (whether linear or non-linear).

QUESTION 3. *Let q, t and d be positive integers such that $q, t, d \geqslant 2$. Does a (q^d, q, t)-perfect hash family S exist such that $|S| < (t-1)d$*?

In other words, are optimal linear perfect hash families also optimal amongst all (not necessarily linear) perfect hash families? Indeed, we may ask whether the following is true:

QUESTION 4. *Let t be a fixed integer, $t \geqslant 2$, and let d be a fixed real number, $d > 1$. Is it true that, for sufficiently large q, a $(\lceil q^d \rceil, q, t)$-perfect hash family S must be such that $|S| \geqslant (t-1)d$*?

REFERENCES

1. N. Alon, Explicit construction of exponential sized families of k-independent sets, *Discrete Math.* **58** (1986), 191–193.
2. N. Alon and M. Naor, Rerandomization, witnesses for Boolean matrix multiplication and construction of perfect hash functions, *Algorithmica* **16** (1996), 434–449.
3. M. Attici, S. S. Magliveras, D. R. Stinson, and W.-D. Wei, Some recursive constructions for perfect hash families, *J. Combin. Designs* **4** (1996), 353–363.
4. S. R. Blackburn, M. Burmester, Y. Desmedt, and P. R. Wild, Efficient multiplicative sharing schemes, *in* "Advances in Cryptology—EUROCRYPT '96" (U. Maurer, Ed.), Lecture Notes in Computer Science, Vol. 1070, pp. 107–118, Springer-Verlag, Berlin, 1996.
5. E. F. Brickell, A problem in broadcast encryption, *presented at* "5th Vermont Summer Workshop on Combinatorics and Graph Theory," June 1991.
6. M. L. Fredman and J. Komlós, On the size of separating systems and families of perfect hash functions, *SIAM J. Algebra Discrete Methods* **5** (1984), 61–68.
7. J. W. P. Hirschfeld, "Projective Geometries over Finite Fields," Clarendon Press, Oxford, 1979.
8. D. R. Hughes and F. C. Piper, "Design Theory," Cambridge University Press, Cambridge, UK, 1985.
9. J. Körner, Fredman–Komlós bound and information theory, *SIAM J. Algebra Discrete Methods* **7** (1986), 560–570.
10. J. Körner and K. Marton, New bounds for perfect hashing via information theory, *European J. Combin* **9** (1988), 523–530.
11. K. Kurosawa and D. R. Stinson, personal communication, 14 June 1996.
12. K. Mehlhorn, "Data Structures and Algorithms," Vol. 1, Springer-Verlag, Berlin, 1984.
13. I. Newman and A. Wigderson, Lower bounds on formula size of Boolean functions using hypergraph entropy, *SIAM J. Discrete Math.* **8** (1995), 536–542.
14. M. E. O'nan, Automorphisms of unitary block designs, *J. Algebra* **20** (1972), 495–511.
15. J. Radhakrishnan, Improved bounds for covering complete uniform hypergraphs, *Inform. Process. Lett.* **41** (1992), 203–207.